DESCRIPTIVE METHODS OF DATA ANALYSIS FOR MARKETING DATA – THEORETICAL AND PRACTICAL CONSIDERATIONS

Manuela Rozalia GABOR
„Petru Maior" University of Tg. Mures, Romania
Nicolae Iorga str. no. 1, Tg. Mures
e-mail: rozalia_gabor@yahoo.com

Management & Marketing
Challenges for Knowledge Society
(2010) Vol. 5, No. 3, pp. 119-134

Abstract. *Marketing has as main objective the guidance of a firm's activities according to current and future needs – of consumers'. This necessarily assumes the existence of a suitable information system, and also the knowledge of some modern analysis, processing and interpretation of the so complex information in the field of marketing.*

The descriptive methods of data analysis represent multidimensional analysis tools that are strong and effective, tools based on which important information can be obtained for market research. The paper comparatively presents some of these methods, respectively: factor analysis, main component analysis, correspondence analysis and canonical analysis.

Keywords: factor analysis, marketing, descriptive methods.

1. Introduction

The data analysis methods were elaborated long time ago, in 1930, H. Hotteling laid the foundation for the *main component analysis* and canonical analysis, thus developing C. Spearman's and K. Pearson's works dating back at the beginning of the century. Also, the main principles of *factor analysis* belong to Spearman (1904), the term as such being introduced much later, in 1931, by Thurstone in psychology. The origins of typological analysis are considered to be two articles published in 1938, of Tyron's, entitled *„A technique for measurement of similitudes with spiritual structures*" and *„General dimensions of individual differences: typological analysis or multiple factor analysis*" among other authors who brought major contributions to typological analysis being: M. Hugues (1970), R. Baechtold (1971), J.F. Canguilhem (1972).

Until the '60s these methods have developed and diversified in versions but however, remained unapproachable in practice as they were requiring a very high amount of calculations. Occurrence of software and PCs enabled the access of patricians to data analysis techniques.

As regards the purposes targeted by data analysis methods, they are various according to specialty authors. Thus, according to Gheorghe Ruxanda, *data analysis has as basic goal the selection of relevant, significant information, that is contained in data, in primary information, this information being used further, for handling some problems specific to data analysis: testing, forecast, interpretation, predictions etc*. According to other author, Carmen Pintilescu, the purpose of data analysis is represented by *distribution analysis of some statistic units based on a set of variables.* G. Saporta and V. Ştefănescu consider that data analysis is *the research of differences and/or similitudes among individuals, considering that two individuals are alike their profiles are close according to various characteristics*, the factor analysis enabling the graph of similitudes and the typological analysis enables their grouping in homogenous categories or that, by means of these methods, relations between characteristics can be described. In the foreign literature, one of the major authors in this field, M. Volle, stated that *„by application of data analysis methods a loss of information is accepted in order to get a better significance*".

Especially the factor analysis methods have represented the basis of developing other methods, for instance the factor analysis on tables of distances and dissimilarities (that has the same purpose as the main component analysis with the difference that initial data is different, knowing only the distances or dissimilarities between individuals and not the variables they describe), the analysis of an Euclidean distance table, in this respect developing the MDSCAL algorithm of J.B. Kruskal that uses ordinal information and the INDSCAL model (*INDividual Differences SCAling*) developed by J.D. Carroll that enables analysis of several distance tables (IDIOSCAL is a second model developed by the same author). Other developed factor methods: PCA of instrumental variables (ACPVI), PCA with orthogonality restriction, PCA with partial co-variances. Among other authors that had major contributions to the development of the descriptive methods of data analysis (especially in the non-metric analysis) the following can be enumerated: F.W. Young, W.S. Torgerson – the latter being related to one of the first software used in data analysis, TORSCA respectively –, J.C. Lingoes, L. Guttman, V.E. McGee.

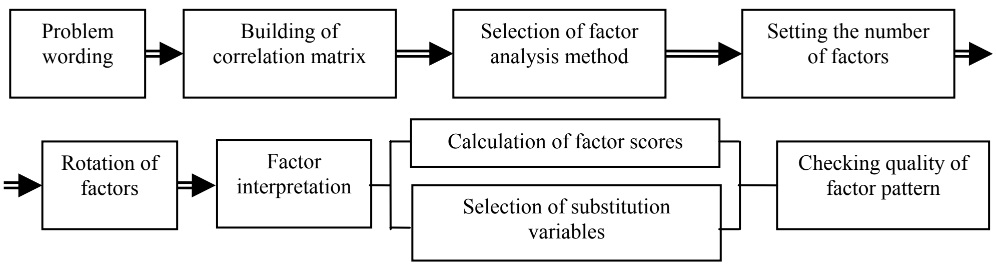
For each method application examples are mentioned for marketing data methods. We mention that, apart these methods in the literature, newer methods are approached within the descriptive methods, multidimensional scaling, conjoint analysis and confirmative structural methods, respectively, Appendix 1 containing the brief presentation of these methods in line with the space localization of cloud of points, the reduced space or total space respectively, when analysis starts and are classified according to the following criteria: visualization, proximity and grouping.

2. Factor analysis

The factor analysis is defined in the literature as being *a method that researches the interdependence relations among several variables whose help, a certain phenomenon is defined, by reducing the amount of information comprised in initial variables and establishment of a smaller set of dimensions (called factors), aiming to a minimum loss of information and focusing on the analysis of the interdependence between them.*

The basic principle in the factor analysis consists in maximization of variance between statistic units concerned and finding the centre lines (components) of cloud of points inertia (variation).

Stages covered in the application of factor analysis methods are illustrated in Figure 1.



Source: Adaptation after Malhorta, N., *Études marketing avec SPSS,* 4e édition, Ed. Pearson Education, France, Paris, 2004, p. 512.

Figure 1. Stages of factor analysis

Each stage mentioned above is important for this method, of which, the *factor rotation* and the *result interpretation* are stages that singularize this method for each type of surveyed problem (economic, social, psychological, marketing etc.) and the literature provides then a wide methodological approach. In the stage of *wording a problem*, using of factor analysis requires that *variables taken into consideration should be measured on a range or a proportional scale*. In the stage of selecting the *analysis method* it relates to the fact that there are two ways of analysis: *the main component analysis* (it will be approached in the following paragraph) and *the common factor analysis*, the latter being used when acknowledgement of common variation becomes a major purpose for analysis (is also called the *main axis factoring*). In order *to set the number of factors* the following procedures can be used: setting the number of factors a priori, factor related variation percentage, slope graph, own values, equal sub-sample analysis or statistic tests.

In fact, the stage of *factor rotation* is only a transformation applied to the factor matrix (allotment) that contains factor loadings. *Statistically*, rotation does not change the value of communality and neither the total percentage of explained variation, but, individually, the rotation method will change the variation percentage explained by each factor. In other words, different rotation methods will be able to result in identification of some different factors. Two types of factor rotations are used, respectively, orthogonal rotation – *when factors obtained are independent* – and inclined rotation – *when factors obtained can be correlated*. The difference between the two types of rotations consists in the factor intersection angle: in case of orthogonal rotation, the centre lines make a square angle meaning that factors are independent, and at inclined rotation, the angle has different values than 90^0, the factors being correlated among them.

For *marketing data*, the *factor interpretation* stage has a major importance to understand the surveyed phenomenon or process, both for quantitative approach and qualitative approach of the factor analysis results. In this stage, apart a very good knowledge of the surveyed marketing aspect, it is required a suitable understanding of the surveyed variables and formulated assumptions concerning relations between variables.

Indicators and statistic notions associated with data factor analysis are shown in Table 1.

Using this method for marketing data is recommended by the fact that, in most market research cases in different situations, the study starts from a multitude of variables of which most of them are correlated (they have common latent elements) enabling and entailing reduction of their number at a workable level.

Table 1

Factor analysis statistic indicators and notions

Indicator or statistic notion used	Description
Correlation matrix	It shows the correlation coefficients of all pairs of model variables.
Comunality	It represents the variation part that a variable has in common with all the other variables included in the analysis model and is, at the same time, the variation proportion explained by common factors.
Own values	It represents the total variation explained by each factor.
Factor loadings	It represents the simple correlations between variables and factors.
Factor loading schedule	It is the graph of primary variables that uses factor loadings as coordinates.
Factor matrix (allotment)	It contains factor loadings of all variables for all selected factors.
Factor scores	There are mixed scores estimated in case of derived factors.
Standard scores	There are score-values related to each individual (each line in data matrix). Standardization is thus carried out so as most scores range between –3 and +3, thus enabling individual ordering.
Factor rotation	It is a change of variable space, whereby factors rotate simultaneously in order to get as many information as more 0 elements in the matrix of factor loading coefficients. (the sum of own values is not affected during this transformation, but rotation will affect its own vectors.)
Measurement of adequacy index (MSA)Kaiser-Meyer-Olkin (KMO)	Index used to evaluate the factor analysis validity, being relevant for high values ranged between 0.5 – 1.
Variation percentage	It represents the part in the factor related total variation.
Residual values	It represents the differences between the noticed correlations (initial), according to the correlation matrix and correlations as a result of estimations carried out based on the factor allotment matrix (after model application). Residual matrix thus resulted can identify errors, discrepancies in model etc.
Statistic tests	The Bartlett roundness test, χ^2 test

Source: Adaptation after Danețiu, T., *Multivariate methods used in computer aided marketing data analysis,* doctoral dissertation, pp. 52-53, Spircu, L., *Data analysis. Economic applications. Course,* ASE, www.ase.ro/bibliotecadigitala, 2004, Malhorta, N., *Études marketing avec SPSS,* 4e édition, Ed. Pearson Education France, Paris, 2004, p. 510.

The factor analysis is in line with some targeted objectives as:

- *Reducing the number of variables* in order to remove redundancies and simplify the study from optional reasons, the purpose being that of keeping only these new variables for further research.

- *Classification of variables*, finding a structure of relations between variables respectively and as a result, interpretation and understanding of obtained factors will be enabled;
- *To identify latent structures* – factors – *explaining correlations within a set of variables;*
- *To identify a new and smaller set of non-correlated variables to replace the original set of correlated variables in analyses to follow (*discriminant analysis or regression analysis*);*

Applications of factor analysis for marketing data are various taking into consideration the multitude and great diversity of variables studied in the field of marketing, thus:

- A low number of statements related to *life style*, obtained following the application of factor analysis, will be able to be used as independent variables in order to explain the differences between loyal customers and the others;
- When identification of *some psychographic profiles of consumers* is required, a set of statements related to life style can be used, these can be analyzed afterwards by means of factor analysis to identify the main psychographic factors on which marketing tools are applied afterwards;
- In *market segmentation* factor analysis can be used to identify the variables on which consumers are grouped. *For instance*, car consumers can be grouped according to the importance they provide to various car related aspects, in segments of users interested in consumption, usefulness, performance, comfort or luxury;
- In *product surveys*, factor analysis enables identification of product attributes that influence consumer choice. Therefore, toothpaste brands can be evaluated as regards cavity protection, tooth whitening, taste, effect of fresh breath and price;
- In *price surveys,* it is used to identify characteristics of price sensitive consumers. For instance, it can be noticed these consumers are concerned by product usefulness, are used to plan shopping and are household focused;
- In *audience surveys*, factor analysis can help understanding the media consumption habits of the target market. Energizing drink consumers can be characterized by a higher TV consumption, as they mainly watch movies and listen to house music;
- For a better *guiding of advertising strategies*, in order to identify those latent factors responsible with grouping the customers in classes.

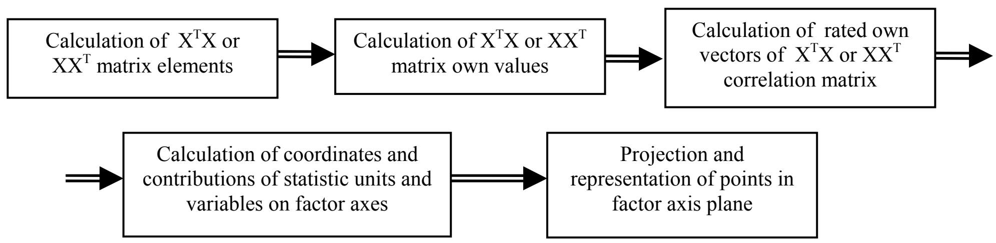
3. Main component analysis (PCA)

The basic principle of this method is to select the lowest number of components to recover as much as possible the total information contained in primary data, these new components expressing new attributes of individuals and built so as they are non-correlated between them, each of these new variables being a linear combination of primary variables. This method provides a graphic visualization of *the map of individuals* in the study according to similarities between them and the *map of variables* according to their correlations.

Although this method is based on the same principle as in the case of factor analysis (in principle it is a linear factor method), the main component analysis differs from it by the way of definition of elements related to initial data table and the calculation way of the distance between points. As a descriptive method of data analysis is applied **only** to quantitative variables and large tables comprising information related to more than 15 individuals and 4 variables. Another characteristic that differentiate it from factor analysis is given by the *way of term transformation* therefore relation (1) is used in the main component analysis, while relation (2) is used in the factor analysis.

$$x''_{ij} = \frac{x_{ij} - \overline{x}_j}{\sqrt{n} * \sigma_j} \quad (1) \qquad\qquad x'_{ij} = \frac{x_{ij} - \overline{x}_j}{\sigma_j} \quad (2)$$

PCA method stages are illustrated in Figure 2.



Source: Pintilescu, C., *Data analysis,* Ed. Junimea, Iaşi, 2003, p. 37.

Figure 2. Stages of main component analysis

The stages presented above are followed by the interpretation of analysis results, G. Saporta and V. Ştefãnescu mention two types of interpretations that should be carried out in case of the main component analysis, *"internal" interpretation* respectively, id est correlations between initial components and variables (represented by the circle of correlations) and *"external" interpretation* between variables and additional individuals, the explanation of results being carried out by means of data that served to their accomplishment.

Two approaches of the main component analysis are described in the literature: *Pearson's geometrical approach* – more laborious but it drives to an assembly of complete results – and *Hotteling's approach,* proposed in 1933 – that suggests criteria (*correlation criterion* and *dispersion criterion*) of directly obtaining the main components but that it has the disadvantage of losing the analysis geometrical dimension of analysis, both widely presented in the quoted Ştefănescu's paper.

In the main component analysis, *to select the number of factor axes the following* criteria will be interpreted:

- *Kaiser criterion* (criterion of improper value) consists in selecting the number of axes for which own values are in line with a value higher than one.
- *Evrard criterion* (criterion of slope or "granularity") based on the graphic representation of own values and tracking a sudden drop of inertia explained by these values.
- *Benzécri criterion* (criterion of cover percentage) assumes selection of that number of axes that explains more than 70% of total variation concerning the cloud of points.
- *Parallel analysis method* (carried out by Horn) is applicable to standard data and assumes generation of random samples, variables that characterize the population being assumed non-correlated two by two.
- *Regression method* is analogue to the parallel analysis but does not assume generation of random samples and performance of the main component analysis at the level of each sample.

Table 2

Statistic indicators and notions related to main component analysis

Indicator or statistic notion used	Description
Own values and vectors	Are associated with the matrix of initial variable correlations. An own value higher than 1, for a component, shows that component has a contribution higher than that of an initial variable, so it is recommended to be selected. Own vectors associated with own values, will represent weights in calculation of the linear combinations concerned.
Scores of main components	Are coordinates of individuals in the new axes, given by selected own vectors. The average of a column of scores is 0.
Schedule of decrease	Provides information concerning own values, but also their decrease rates.
Mahalanobis distance	They are used to measure the distance between an individual and the centre of gravity of the cloud of points.

Circle of correlations	Has coordinates made of correlation factors between initial variables and main components related to the factors retained.
Kaiser-Guttman rule	Is used to set the number of main components: there are so many components as own values are higher than 1. (However, the final number of components will be set according to the actual interpretation they receive.)
Inertia criterion	Is used to get the main component and has the advantage of a geometrical approach and is much more complex compared to the correlation criterion and the dispersion criterion suggested by Hotteling.
Global quality of the first main component	For this is used the total inertia decomposition in inertia explained by the first axis and residual inertia of the cloud of points around the first factor axis.
Main components	Are abstract vector variables defined as some linear combinations of original variables and have two basic properties: are not correlated two by two and the first component is a normalized linear combination whose variance is maximum.
Loading coefficients	Are the correlation coefficients between the original variables and scores? They express the importance of each original variable in explaining each new component.

Source: Adaptation after Spircu, L., *Data analysis*, *Economic applications.* Course, www.ase.ro/bibliotecadigitala, 2004, Spircu, L., Calciu, M., Spircu, T., *Marketing data analysis*, Ed. ALL, Bucharest, 1994, pp. 92-98, Ştefănescu, V., *Data analysis – case studies*, Ed. ASE, Bucharest, 2000, pp. 11-18.

In order to select the number of main components standard linear combinations are used that have as a starting point instead of the R matrix of correlations, the covariance sum matrix, and consists in selecting that standard linear combination with the highest variance. *Apart the factor analysis* – wherein variations of X variables by means of linear transformations of a fixed number, limited by factors called ”hidden”, latent – the main component analysis is a searching linear combinations between variables being set in order based on own values of covariance matrix.

As practical examples for marketing data of the main component analysis we can mention:

- Analysis of summer vacation distribution starting from their distribution, in a certain period of time, according to the socio – professional categories that requested such vacations and based on the manner of accommodation selected by tourists;
- The main component analysis applied for the survey of brands related to some commodities or some durables;
- The study on the quality of some food brands existing on the market and their breakdown according to quality;

• The main component analysis is used to handle the various dimensional scaling problems in the field of marketing;

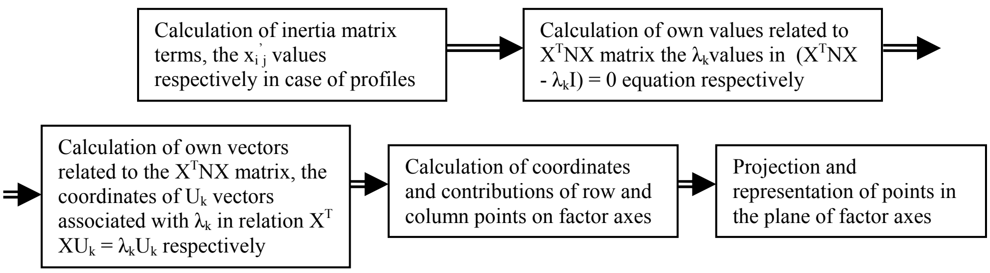
• Evaluation of a bank's services by its clients, starting from the breakdown of banks by clients according to their importance and 15 bank related qualities, the key factors resulted may be: traditional services, convenience, visibility and competence.

4. Correspondence analysis

This method has been suggested by J.P. Benzécri in France for studying not only for testing the independence between variables as well as for describing the association (also called correspondence) between two qualitative variables. This method emphasizes the relations between methods, and on the other hand it provides the opportunity of a relatively easy readable graphics.

Mathematically, we can consider the correspondence analysis *as a main component analysis with a special metric, $\chi 2$ metric*, or as a version of *canonical analysis*. G. Saporta and V. Ştefãnescu, of the two versions develop in the quoted paper, the *aspect of canonical analysis of correspondence analysis* as it presents the benefit of complying with symmetry between the two variables and to generalize without trouble the correspondence analysis to several qualitative variables. For the other version, compared to the other factor analysis methods, using the χ^2 distance between two points, the main component analysis using for instance the Euclidean distance is emphasized as feature.

The stages of correspondence factor analysis method are briefly represented in Figure 3.



Source: Adaptation after Pintilescu, C*., Data analysis,* Ed. Junimea, Iaşi, 2003, p. 77.

Figure 3. Stages of correspondence factor analysis

The initial data for this method is comprised in the *contingency table* and, as a comment, the method can be also used for the correspondence analysis between *numeric variables*, not only nominal, on condition they have positive values, and their sum on row and column is different than zero in order to apply the correspondence analysis. Also, the correspondence factor analysis can be also carried out by reversing rows with columns so as they do not change the results at all, apart the main component analysis where their reversion would have a back action.

In order *to select the number of factor axis*, *empirical criteria is used*, the most common being that of *selecting a curve in the diagram of own values also* used in the main component analysis.

Table 3 briefly shows the main statistic notions and indicators used in correspondence analysis.

Table 3

Indicators and statistic notions associated with correspondence analysis

Indicator or statistic notion used	Description
Total inertia	Or sum of own values, calculated as ratio between χ^2 and *n*
Tables of row profiles and column profiles	Have as terms relative frequencies partly calculated as a ratio of absolute partial frequencies and the marginal ones for each *i* method of variable
Variable independence	Between *X* and *Y* variables there is not any relation if, for observed data values of partial relative frequencies are met so as $f_{ij} = f_i \times f_j$
χ^2 *distance*	The adjacency between two lines or two columns is the distance between their profiles.
Plain value	It is the first own value λ_1 and is not interpreted in the correspondence factor analysis.
Testing of hypotheses	χ^2 test

L. Spircu, M. Calciu and T. Spircu give this method as an example in case when a brewery that carries out a sampling based research in order to get the data required to elaborate its strategy of quality diversification. The sample has been made of respondents of various ages and occupations that were questionned in connection with the favorite beer and the favorite newspaper, *characteristic ”favorite beer brand” being against all the others*, thus applying *the general case of factor analysis* proposed by Benzécri (correspondence analysis having another application version, for two characteristics).

Another example of using the correspondence analysis is illustrated by G. Saporta and V. Ştefănescu in the quoted paper related to analysis of cigarette brands, the survey being carried out in France and comprises the correspondence analysis between four levels of preferences, four levels of price, four levels of cigarette qualities and three lots of interviewed smokers.

For the marketing data from samplings that used questions with multiple choices in the questionnaire, a version of the correspondence analysis is recommended, multiple correspondence analyses respectively, considered by some authors as being a main component analysis for qualitative variables, the result interpretation being carried out in case of correspondence analysis and main component analysis.

We should mention that in the factor analysis of correspondences, apart the main component analysis, analysis of indicators calculated for the points rows and columns is indispensable and as in the main component analysis, correspondence analysis finally results in emphasizing some new factors that are represented by own vectors of the products related to the two profile matrices. Correspondence analysis is a method recommended in the study of qualitative variables and one of the main major advantages is given by the possibility of representing simultaneously the rows and columns of a contingency table.

5. Canonical analysis

Canonical analysis as data descriptive analysis method belongs to Hotteling, that, in 1936 extended analysis of correlations between quantitative variables to analysis of correlations between two sets of quantitative variables (and finding that linear combination that represents the set of variables optimally) and not only two variables based on the Pearson correlation coefficient.

When canonical analysis is used the following objectives are mainly targeted:

- Determination of a maximum correlation between p explained variables and q explanatory variables;
- Determination of weigths for each assembly of explained variables and explanatory variables so as the sum of weights thus obtained is correlated as much as possible;
- Obtaining other linear functions by maximizing the remained dependencies, non-correlated with the assembly of previously obtained weights.

When n individuals are described by two aggregates of variables $(p$ and $q)$ examination of the existing correlations between the two aggregates is requested in order to know if they measure the same properties or not. As examples we can consider: two groups of marks at literary and scientific subjects, results of medical analyses performed at two different laboratories.

In the literature a generalized version of canonical analysis is mentioned. This version developed by J.D. Carroll has as principle searching an auxiliary variable z that belongs to W_i subspaces so as $\Sigma R^2(z, X_i)$ is maximum, the method mainly being a main component analysis for groups of variables.

Table 4

Statistic indicators and notions related to canonical analysis

Indicator or statistic notion used	Description
Own values	Squares of canonical correlation coefficients between canonical variables are given.
Statistic tests	Bartlett test
"Forecast potentials"	These two subspaces W_1 and W_2 id est aggregates of variables that we can build by linear combinations of the variables of the two groups with $W_1 = \{x/x = X_1a\}$and $W_2 = \{y/y = X_2b\}$. If these two spaces are identical with, it proves that we can do only one of the two aggregates of variables, as they have then the same description strength; if they are orthogonal, it means the two aggregates of variables show different phenomena.

Source: Adaptation after G. Saporta, V. Ştefănescu, *Data analysis and computer science,* Ed. Economică, Bucharest, 1996, p. 96.

Applications of canonical analysis for marketing data:

- Consumers' loyalty for a product can be expressed by some variables: probability of doing a shopping, the space of time between two shoppings and the amount purchased at one time, that can be the set of dependent variables used in canonical analysis. Consumers' loyalty is driven by a set of variables consisting of product attributes (price, color, packaging, taste etc.) that can be the set of independent variables.
- Evaluation of relation between performances of a brand (measured by sales, market share, sales growth, profit etc.) and variables of marketing mix (price, promotion, distribution, product);
- Evaluation of relation between attitude concerning breakfast as important meal in nutrition and buying of products intended to consumption at breakfast.

Although direct applications of canonical analysis are not many, however it remains a basic method, because its procedure (searching of couples with maximum correlation variables) is found in other methods as correspondence analysis or discriminatory analysis methods. Canonical analysis shows some methodological *disadvantages* for which researchers uses much less, respectively: the method aims maximization of correlation in linear combinations and not of extract variance; canonical correlation reflects variation divided between linear combinations and not in extract variance and therefore a strong canonical correlation may result, the extract variance can be very weak.

6. Conclusions

In marketing, an individual – consumer, buyer etc. – is defined by more than a variable and the other statistic methods (for instance correlations) enable analysis of each variable, but separately, while data descriptive analysis – and especially the main component analysis – enables approaching of data multidimensional character, variables that defines an individual.

One of the disadvantages – in our opinion – of using these mehods, is the fact that, for some of them (for instance, *factor analysis*) there is not a statistic test of testing significance and having the possibility of concluding that results obtained are not accidental but reflect something significant. A solution to this problem can be dividing the sample randomly composed of two sub-samples and application of each of them to the factor analysis. If the same factors result in each of the two sub-samples than we conclude that results are not due to chance.

Therefore we may conclude that, the descriptive methods of data analysis can be used successfully in the following cases:

- To *identify dimensions or basic factors those explain correlations between many variables*. For instance, an assembly of items related to life style can be used to measure the consumers' psychographic profiles;
- To *identify a new and smaller assembly of non-correlated variables, in order to replace the first assembly of variables correlated in a multivariate analysis* (regression analysis or discriminant analysis). For instance, the identified psychographic factors can be used as independent variables to explain differences between loyal consumers and unreliable consumers;
- To *identify a smaller assembly of determinant variables* starting with a bigger assembly so as to be applied to a multivariate analysis. For instance, some items on the original life style in correlation with the identified factors can be used as an independent variable in order to explain the differences between loyal consumers and unreliable consumers.

Studying the many marketing variables that in most cases are correlated between them by means of data descriptive analysis is very important and represents a useful information for the management and marketing of a company.

References

Aaker, D.A., Kumar, V., Day G.S. (1998), *Marketing research,* Sixth edition, John Wiley & Sons, New York, p. 573

Benzécri, J.P., Benzécri, F. (1980), *L'analyse des donnéesi,* Dunod, Paris, pp. 40- 42

Cătoiu, I. (coord.) (2002), *Marketing research,* Ed. Uranus, Bucharest, p. 534, pp. 539-540

Dănețiu, T. (2005), *Multivariate methods used in computer aided marketing data analysis.* Doctoral dissertation, ASE, pp. 48-50, 52

Gabor, M.R. (2009), "Typological analysis method of marketing data"*, Management & Marketing,* no. 4 (16)/2009, Bucharest, pp. 125-132

Gabor, M.R. (2009), "Analysis of marketing data – evolutions of topic approach", *Analele Universității din Oradea, Economics Series,* tome XIX, vol. IV/ 2009, ISSN 1582-5450, CD-ROM

Gabor, M.R. (2009), "Statistics test used in data analysis", *Management and Economic Engineering*, Volume 8, no. 3(33)/2009, pp. 185-194

Gauthy Sinéchal, M., Vandercammen, M. (2005), *Études de marchés–methodes et outils –* 2e édition, De Boeck & Larcier, Buxelles, pp. 340-341

Giannelloni, J.L., Vernette, E. (2001), *Etudes de marche* – 2e edition, Vuibert, Paris, p. 382

Green, P.E., Tull, D.S. (1974), *Recherche et decisions en marketing,* PUG, pp. 359-360

Malhorta, N. (2004), *Études marketing avec SPSS,* 4e édition, Pearson Education, Paris, p. 507, 510, 512

Pintilescu, C. (2003), *Data analysis,* Junimea, Iaşi, p. 8, 10, 24, 37, 77

Ruxanda, Gh. (2001), *Data analysis,* ASE, Bucharest, 2001, p. 81, p. 418, pp. 507-509, 511

Saporta, G., Ştefănescu, V. (1996), *Data analysis and computer science,* Ed. Economică, Bucureşti, 1996, pp. 76-80, p. 87, 92, 96, 100, 105, 125, 238, 287

Smajda, A. (1988), *Segmenter ses marches – applications pratique de technique de segmentation dans le marketing*, PPR, 1988, p. 45

Spircu, L., Calciu, M., Spircu, T. (1994), *Marketing data analysis,* Ed. ALL, Bucureşti, 1994, pp. 92-99, 102

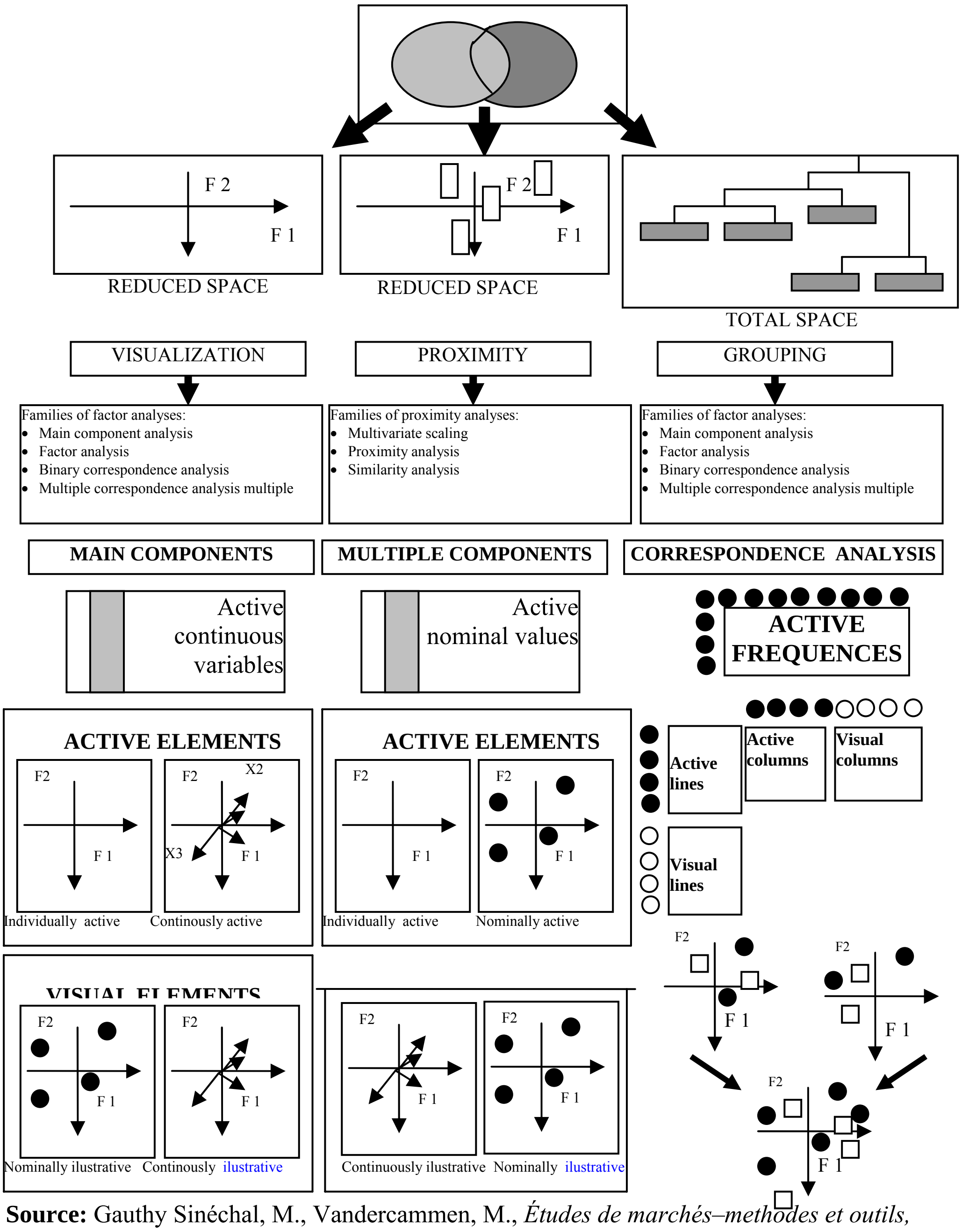
Ştefănescu, V. (2000), *Data analysis – case studies*, Ed. ASE, Bucureşti, pp. 11-18

Volle, M. (1997), *Analyse de données* - 4e edition, Economica, Paris

Appendix 1

Descriptive methods of data analysis

Starting from point cloud in space



Source: Gauthy Sinéchal, M., Vandercammen, M., *Études de marchés–methodes et outils,* 2e édition, Ed. De Boeck & Larcier, Buxelles, 2005, pp. 340-341.